DOI: 10.1007/s10535-018-0824-1 BIOLOGIA PLANTARUM 62 (4): 835-839, 2018

BRIEF COMMUNICATION

Alternative splicing at the *o2Italian* locus in maize: one mutation, two proteins

I. MASCHERETTI, A. GRASSO, A. VIOTTI, and M. LAURIA*

Istituto di Biologia e Biotecnologia Agraria, CNR, I-20133 Milano, Italy

Abstract

Opaque2 is a basic leucine-zipper transcription factor that regulates the expression of several genes involved in the development of maize endosperm. Among the different mutations so far characterized, the *o2Italian* (*o2It*) allele has a genomic deletion of 10 bp that interrupts the open reading frame by a premature stop codon. Curiously, the *o2It* allele still encodes a long o2 isoform (o2It-L) that is recognized by an O2 carboxy-terminal antibody, evidencing that this protein is similar to the wild type isoform. To clarify this phenomenon, we characterized the transcripts generated from the *o2It* allele and found two transcripts that are produced in a 1:1 ratio. The first transcript carried the 10 bp deletion and matched the genomic sequence of the *o2It* allele, whereas the second one carried a 15 bp deletion as a result of a 3' alternative splicing site. The lack of five additional nucleotides restored the correct reading frame and explained the nature of the o2It-L polypeptide. The production of two distinct transcripts from the *o2It* allele was probably due to the recruitment of different splicing factors as suggested by *in silico* analysis of the *o2It*-mutated region. This finding evidences how the recruitment of splicing factors is tightly linked to nucleotide motives that should be present in correct neighboring contexts.

Additional key words: basic leucine-zipper transcription factor, endosperm, open reading frame, *Zea mays*.

The Opaque2 (O2) transcriptional activator is a basic leucine-zipper protein of maize (*Zea mays*) specifically expressed in the endosperm (Schmidt *et al*. 1987, 1990). The O2 protein regulates multiple metabolic pathways during seed development (Li *et al*. 2015), and its best characterized targets are the 22-kD α-zeins: a subfamily of maize seed storage proteins (Song *et al*. 2001). These proteins are poor in lysine content and, consequently, they negatively affect the nutritional value of maize seed (Pirona *et al*. 2005). Because mutations at the *O2* locus determine a reduction of the 22-kD α-zeins with a consequent increase of lysine rich proteins in the seed, the *o2* mutation has been the subject of intensive studies for several decades (Pirona *et al*. 2005). In maize, the *o2* alleles so far identified are recessive mutations and can be subdivided in two categories. The first includes alleles that produce mutated o2 polypeptides, such as the *o2T*, *o2It*, *o2-52*, and *o2-676*, whereas the second includes null-transcript alleles named *o2R* (Schmidt *et al*. 1987, Motto *et al*. 1988, Bernard *et al*. 1994, Hoschek *et al*. 1996, Lazzari *et al*. 2002, Gavazzi *et al*. 2007). By using antibodies that recognize the N- and C-terminal part of the O2 protein (O2-NT and O2-CT), it was found that four maize lines carrying the *o2Italian (o2It)* allele produced two polypeptide isoforms (Gavazzi *et al*. 2007). The first isoform migrates slightly faster than the wild type and are named o2It-Long (o2It-L), whereas the second isoform has a relative molecular mass of about 47 kDa and was named o2It-Short (o2It-S). Both the isoforms were recognized by the O2-NT antibody, whereas only the o2It-L isoform is recognized by the O2-CT antibody. Molecular characterization of the *o2It* allele revealed the presence of 10 bp deletion at the beginning of the fifth exon, which introduces a premature

Submitted 26 January 2018, *last revision* 28 April 2018, *accepted* 23 May 2018.

Abbreviations: DAP - days after pollination; ESE - exonic splicing enhancer; O2 - Opaque2; ORF - open reading frame; PC - primer combination..

Acknowledgments: The authors would like to thank Dr. Corrado Viotti for critical reading and editing the manuscript. This work was supported by a special grant from the Italian CNR Flagship project EPIGEN.

* Corresponding author; e-mail: lauria@ibba.cnr.it

stop codon (Lazzari *et al*. 2002). Therefore, the presence of a nearly full-length protein, the o2It-L isoform, in maize lines carrying the *o2It* allele was incoherent. Most puzzling was that the o2It-L isoform is detected by the O2-CT antibody, which is specific for the last 170 amino acids of the O2 protein. This finding suggests that O2 and o2It-L must be nearly identical in term of amino acid sequence. To explain the presence of the o2It-L isoform, it was proposed a mechanism based on transfer RNAs that suppresses the non-sense codons and promotes a frameshift of the transcript (Lazzari *et al*. 2002), a mechanism already observed in other model organisms (Baum and Beier, 1998, Golovko *et al*. 2002, Lecointe *et al*. 2002). In this study, however, we investigated the possibility that the two o2 isoforms originate from two different transcripts.

The maize lines NYR*O2* (N+), NYR*o2It* (N*o2*), and Rossman-*o2R* (R*o2*) were grown either in the field or greenhouse and reproduced by self-pollination. Immature seeds were harvested at 18 - 19 d after pollination (DAP), frozen in liquid nitrogen and stored at -80 °C.

Genomic DNAs and total RNAs were extracted from dissected endosperms by using the cetyltrimethyl-ammonium bromide (CTAB) protocol and the *TRIzol*® reagent (*Ambion*, Carlsbad, USA), following the manufacturer's instructions. DNA and RNA quality/quantity was determined by *Nanodrop* (*Thermo Scientific*, Waltham, USA). For cDNA synthesis, 10 µg of RNA for each sample was treated with *TURBO* DNA-free DNase (*Ambion*) to remove genomic DNA. Then, a total of 2 µg of RNA DNase-treated was used for the first-strand cDNA synthesis with *SuperScript II* (*Invitrogen*, Carlsbad, USA) and 500 ng of oligo(dT)18 primer at 42 °C for 1 h in a final volume of 20 mm^3, following the manufacturer's instructions. For 3'-RACE analysis the *FirstChoice® RLM RACE* kit (*Ambion*) was used, following the manufacturer's instructions.

PCR analysis of cDNA and genomic DNA was performed as follows with the exception of annealing temperature that is specific for each primer combination (PC) used: 95 °C for 5 min; 95 °C for 30 s, 58 - 60 °C for 30 s, 72 °C for 45 s for 36 - 38 cycles; then 72 °C for 4 min. The specific primer combinations (PCs) used to clone the entire *O2* and *o2It* sequence were designed based on the *O2* sequence of B73 line (GRMZM2G015534) (http://www.maizegdb.org/). In detail, F1 (GAAAGAAGAGCATCCAAGCGT) and R1 (ATCAACCGGAATTACCTTTTCCTC) PC were used for the amplification of the 5'-end region. The F2 (GAACGGCGAAGATCCAATCAA) and R2 (CTTAGC TCTTAGGGTCTCCATGT) PC were used for the amplification of the internal region and as well as for the characterization of the *o2* transcripts spanning the fourth and fifth exons. The F3 (GCACAGCTAAAAGCC GAGAAT) and R3 (AATACATGTCCATGTGTATGG) PC was used for the amplification of the 3'-end region.

For Western blot analysis total proteins were extracted from dissected immature endosperms (18 - 19 DAP) and ground in liquid nitrogen. The powder was mixed with extraction buffer [200 mM Tris-HCl pH 6.8, 40 % (v/v) glycerol, 8 % (m/v) sodium dodecyl sulphate, and 20 % (v/v) β-mercapto-ethanol] and incubated at 65 °C for 5 min. The solution was placed on ice for 5 min and centrifuged at 12000 *g* at room temperature for 10 min. The protein extracts were then separated on 12 % (m/v) SDS-PAGE and transferred to a nitrocellulose membrane (*Perkin Elmer*, Norwalk, USA) at 300 mA and temperature of 10 °C for 60 min. The membrane was blocked by incubation with 5 % (m/v) defatted milk in TBST [50 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 0.1 % (v/v) *Tween 20*] for 1 h. Then, the membrane was incubated at room temperature for 1 h with the primary antibody (1:2000), dissolved in 5 % defatted milk in TBST. After washing three times for 10 min with TBST, the secondary antibody (1:16000) was added in 5 % defatted milk in TBST solution and incubated for 1 h at room temperature, followed by three additional washing steps with TBST for 10 min. For detection, the *SuperSignal West Pico* chemiluminescent substrate kit was used (*Thermo Scientific*), following the manufacturer's instructions. The membranes were exposed and analyzed with the *ChemiDoc MP* system (*Bio-Rad*, Hercules, USA).

Genomic DNA bisulfite treatment was performed with *EpiTect Bisulfite* kit (Qiagen, Valencia, USA) following the manufacturer's instructions. Primers for the upper strand were designed using the *KismethWeb*-based tool (http://katahdin.mssm.edu/kismeth/revpage.pl; Gruntman *et al*. 2008). The bisulfite-treated DNA (1 mm^3) was used for PCR analysis as follows: 95 °C for 5 min; 95 °C for 30 s, 45 °C for 1 min, 72 °C for 1 min, 38 cycles; 72 °C for 4 min. The genomic region of *o2It* spanning the fourth and fifth exons after bisulfite treatment was amplified with the FB (GGAGAA GTAGAGATTYTGGGGTT) and RB (CACCARCTT AAAAACCCTTTTTTT) PCs.

PCR amplicons generated from cDNA, genomic DNA, and bisulfite treated DNA were cloned in *pGEM-T Easy* vector system (*Promega*, Madison, USA), following the manufacturer's instructions. Clones were sequenced at the *GATC-Biotech* (Cologne, Germany). For bisulfite analysis, at least 10 independent clones were sequenced. The *Clustal Omega* tool (http://www.ebi.ac.uk/ Tools/msa/clustalo/) and the *KismethWeb*-based tool were used for sequencing.

We confirmed the presence of the o2It-L and o2It-S isoforms in the material currently available in our laboratory by immunoblot analysis of total protein extracted from immature endosperms of the N*o2* and N+ (wild type) lines. The analysis was conducted using the O2-NT and O2-CT antibodies that recognize the N-terminal and C-terminal parts of the O2 protein,

respectively (Gavazzi *et al*. 2007). Accordingly, to previous data, we found that in the N+ line both antibodies detected a single protein with a relative molecular mass of about 68 kDa. On the contrary, in the N*o2* line, two proteins, corresponding to the o2It-L and o2It-S isoforms, were detected with the N-terminal antibody (Fig. 1*A*). In order to identify transcripts able to explain the presence of the two different o2 isoforms, we amplified, cloned, and extensively sequenced cDNA clones that overlapped the *o2It* 10 bp deletion. Sequence alignment confirmed the presence of two *o2It* transcript variants that were produced in a nearly 1:1 ratio (Fig. 1*B*). The first transcript contained the already known 10 bp deletion (Lazzari *et al*. 2002), whereas the second one lacked five additional nucleotides that generated a deletion of 15 bp. This newly identified deletion was totally included within the exon and, therefore, it restored the coding frame, permitting the production of an o2 protein that was only five amino acids shorter than the wild type O2. To confirm that the two transcript variants were correctly processed until the stop codon by the RNA polymerase complex, a 3'-RACE analysis was performed. We found that both transcripts were polyadenylated and, therefore, able to be translated (Fig. 1*B*).

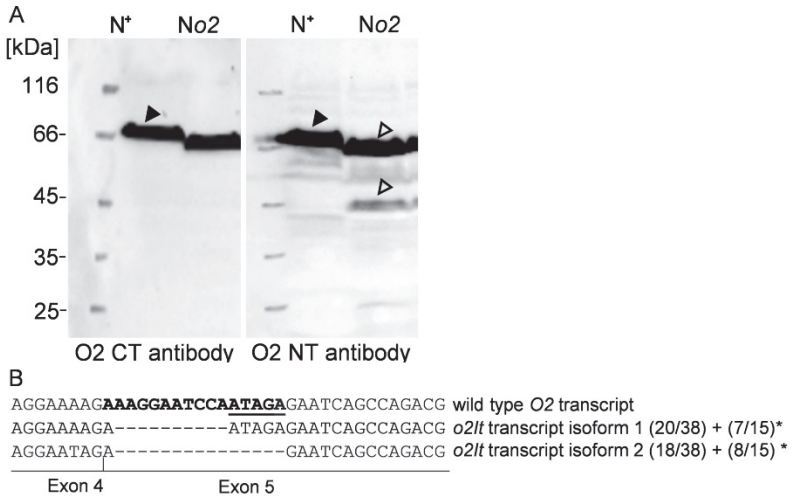

Fig. 1. *A* - Immunoblotting analysis of endosperm proteins extract from N+ and N*o2* lines with the O2-NT and O2-CT antibodies. The *solid arrow* indicates the wild type O2 isoform, whereas the *empty arrow* indicates the o2It-L and o2It-S isoforms. *B* - Sequence alignments between *O2* and *o2It* cDNA sequences. The *o2It* isoform 1 encodes for the o2It-S; the *o2It* isoform 2 encodes for the o2It-L. For simplicity, only the nucleotide sequence alignment spanning the mutated sequence is shown. The 10 bp deletion is *in bold*, whereas the 5 bp polymorphism between the two *o2It* transcripts is *in bold and underlined*. The fraction inside parentheses indicates the number of sequences obtained for each isoform on the total number of sequences analyzed. * indicates the number of sequences obtained in RACE experiments.

The 10 bp deletion that characterizes the *o2It* allele was originally identified by sequencing the cDNA synthesized from the N*o2* RNA (Lazzari *et al*. 2002). Therefore, the presence of two transcripts raised the question about the true nature of the *o2It* mutation. To clarify this point, we cloned and sequenced the entire *O2* genomic region of the N*o2* line by using a PCR approach. Moreover, the same approach was used to obtain the *O2* and *o2* full-length sequences of the N+ and R*o2* null transcript mutant lines, respectively. The nucleotide alignment of the four genomic sequences, including the sequence of the reference line B73, confirmed that the 10 bp deletion present at the beginning of the fifth exon is the only DNA polymorphism that can explain the *o2It* mutation (Fig. 1 Suppl.). In this context, we can conclude that the *o2It* allele generated two mRNA variants that gave rise two o2 polypeptides, whose sizes were in line with the two bands detected by Western-blot (Fig. 1*A*). Notably, the o2It-S isoform was less abundant than o2It-L, suggesting that the former could undergo some kind of quality control process. In plants, the presence of premature stop codons can trigger mechanisms of surveillance/control that either avoid the production of truncated proteins or that degrade them. The similar abundance of the o2It-S and o2It-L transcripts observed in the N*o2* endosperm would suggest that the low abundance of o2It-S depended on protein quality control mechanisms and the ubiquitin-proteasome degradation system (Defenouillère and Fromont Racine 2017).

The simplest explanation for the presence of two *o2It* transcripts was the occurrence of an alternative splicing event. During mRNAs maturation, the exons must be correctly identified and joined together by the coordinated action of small nuclear RNAs and up to 100 accessory proteins (Cartegni *et al*. 2002, Faustino and Cooper 2003). Because *o2It* deletion occurred near to the intron-exon boundary, the canonical 3' splicing site of the fifth exon may be partially replaced by a cryptic and

normally unused 3' splicing site. Beside classical splicing signals, however, additional *cis*-acting elements within the coding region are important for the identification of the correct splicing-site (Cartegni *et al*. 2002, Lopato *et al*. 2002, Xiao *et al*. 2007). The exonic splicing enhancers (ESEs) are auxiliary *cis*-elements frequently found in most exons, including constitutive ones: the exons that are always included in the mature mRNAs, even in different mRNA isoforms (Cartegni *et al*. 2002). The ESEs are recognized by a group of serine/arginine-rich (SR) proteins, which support the splicing machinery in recognizing splice sites (Cartegni *et al*. 2002). Because the basic splicing mechanism appears to be conserved, at least to some extent, between plants and animals (Xiao *et al*. 2007), we performed an *in silico* analysis with a tool originally developed for the animal model system (Cartegni *et al*. 2003) to understand if the *o2It* mutation may affect the recruitment of specific splicing factors. By comparing the mutated region of the *o2It* allele with the same region present in the N+ and B73 wild lines, we found that at least one splicing factor (SRp40) is differentially recruited on the *o2It* mutated region, because of depletion of two binding sites (Fig. 2). Thus, one possibility is that the lack of SRp40, homologs of which in *Arabidopsis* are AtSR33/AtSCL33, AtSCL30a, AtSCL30 and AtSCL38, and AtSRp34a and AtSRp34/SR1 (Wang and Brendel 2004), could cause alternative splicing in about 50 % of the *o2It* mature transcripts.

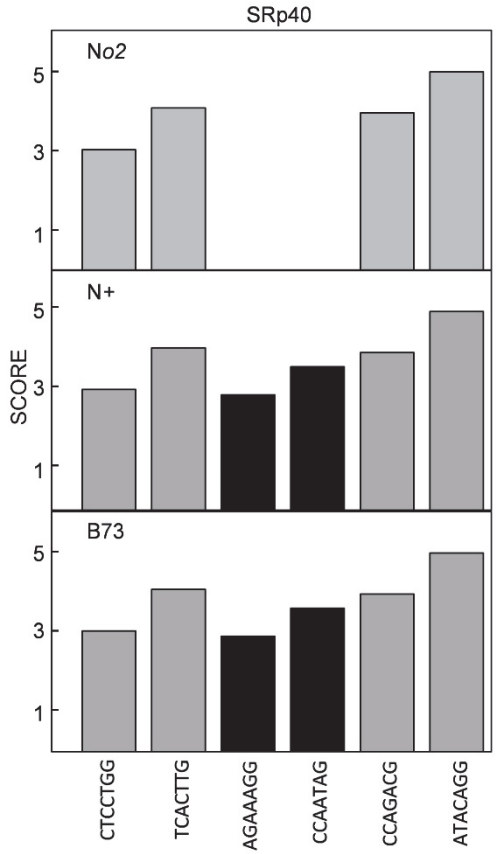

Fig. 2. Exonic splicing enhancer analysis of the *O2* and *o2It* alleles. The sequence-specific binding sites present in the *o2* mutated regions are reported along the X-axis. The height of histograms represents the motif scores. The black histograms indicate that the binding sites are depleted in the *o2It* allele.

It is interesting to underline that the splicing mechanism can also be influenced by epigenetic modifications, which can alter chromatin structure and the elongation rate of RNA polymerase (Luco *et al*. 2010, 2011, Maor *et al*. 2015, Yearim *et al*. 2015). In particular, DNA methylation is frequently present within the gene body, and it is enriched in exons and at exon-intron junctions, where it overlaps with nucleosomal DNA (Chodavarapu *et al.* 2010). Therefore, we asked whether the 10 bp deletion may alter nucleosomes positioning and consequently the deposition of DNA methylation. To answer this question, we analyzed the DNA methylation profile of *O2* and *o2It* alleles in N+ and in N*o2* lines, respectively, with particular attention to the region in which the DNA polymorphism is present. As reported in Table 1 Suppl., bisulfite sequencing analysis showed no differences of DNA methylation in all three C-contexts (CG, CHG and CHH, where H could be A, T, or C), between the N+ and N*o2* lines, thus, suggesting that this epigenetic mark is not involved in the phenomenon under study.

In conclusion, our analysis revealed that the presence of the o2It-S and o2It-L isoforms are the result of two distinct transcripts arising from an incorrect splicing event. The *o2It* mutation produces a non-functional protein since the premature stop codon determines the formation of a transcriptional factor that lacks of its basic leucine zipper region (Ciceri *et al*. 2000). Curiously, the *o2It* mutation also promoted the use of a 3' alternative splicing site that by chance generates a transcript that encodes for a protein nearly identical to the wild type one. Thus, our study highlights how splicing regulation wa tightly linked to nucleotide motives that should be present in correct neighboring contexts. Given the importance of alternative splicing in both plant stress adaptation and development (Dubrovina *et al*. 2013, Shang *et al*. 2017), the *o2It* allele might be an interesting tool to further clarify mechanisms of reciprocal recognition/interaction between nucleotide motives and the splicing protein complexes.

References

Baum, M., Beier, H.: Wheat cytoplasmic arginine tRNA isoacceptor with a U*CG anticodon is an efficient UGA suppressor *in vitro*. - Nucl. Acids Res. **26**: 1390-1395, 1998.

Bernard, L., Ciceri, P., Viotti, A.: Molecular analysis of wild-type and mutant alleles at the *Opaque-2* regulatory locus reveals different mutations and types of *O2* products. - Plant mol. Biol. **24**: 949-959, 1994.

Cartegni, L., Chew, S.L., Krainer, A.R.: Listening to silence and understanding nonsense: exonic mutations that affect splicing. - Natur. Rev. **3**: 285-298, 2002.

Cartegni, L., Wang. J., Zhu, Z., Zhang, M.Q., Krainer, A.R.: ESEfinder: a web resource to identify exonic splicing enhancers. - Nucl. Acids Res. **31**: 3568-3571, 2003.

Chodavarapu, R.K., Feng, S., Bernatavichute, Y.V., Chen, P.Y., Stroud, H., Yu, Y., Hetzel, J., Kuo, F., Kim, J., Cokus, S.J., Casero, D., Bernal, M., Huijser, P., Clark, A.T., Krämer, U., Merchant, S.S., Zhang, X., Jacobsen, S.E., Pellegrini, M.: Relationship between nucleosome positioning and DNA methylation. - Nature **446**: 388392, 2010.

Ciceri, P., Castelli, S., Lauria, M., Lazzari, B., Genga, A., Bernard, L., Sturaro, M., Viotti, A.: Specific combinations of zein genes and genetic backgrounds influence the transcription of the heavy-chain zein genes in maize *opaque-2* endosperms. - Plant Physiol. **124**: 451-460, 2000.

Defenouillère, Q., Fromont Racine, M.: The ribosome bound quality control complex: from aberrant peptide clearance to proteostasis maintenance. - Curr. Genet. **63**: 997-1005, 2017.

Dubrovina, A.S., Kiselev, K.V., Zhuravlev, Y.N.: The role of canonical and noncanonical pre-mRNA splicing in plant stress responses. - BioMed Res. Int. **2013**: 1-14, 2013.

Faustino, N.A., Cooper, T.A.: Pre-mRNA splicing and human disease. - Genes Develop. **17**: 419-437, 2003.

Gavazzi, F., Lazzari, B., Ciceri, P., Gianazza, E., Viotti, A.: Wild-type Opaque2 and defective opaque2 polypeptides form complexes in maize endosperm cells and bind the Opaque2-zein target site. - Plant Physiol. **145**: 933-945, 2007.

Golovko, A., Sitbon, F., Tillberg, E., Nicander, B.: Identification of a tRNA isopentenyltransferase gene from *Arabidopsis thaliana*. - Plant mol. Biol. **49**: 161-169, 2002.

Gruntman, E., Qi, Y., Slotkin, R.K., Roeder, T., Martienssen, R.A., Sachidanandam, R.: Kismeth: analyzer of plant methylation states through bisulfite sequencing. - BMC Bioinform. **9**: 371, 2008.

Hoschek, G., Kelly, A.R., Schmidt, R.J.: Characterization of the molecular defect in a null allele of the opaque-2 locus of maize. - Plant Mol. Biol. **32**: 1159-1161, 1996.

Lazzari, B., Cosentino,. C, Viotti A.: Gene products and structure analysis of wild-type and mutant alleles at the *Opaque-2* locus of *Zea mays*. - Maydica **47**: 253-265, 2002.

Lecointe, F., Namy, O., Hatin, I., Simos, G., Rousset, J.P., Grosjean H.: Lack of pseudouridine 38/39 in the anticodon arm of yeast cytoplasmic tRNA decreases *in vivo* recoding efficiency. - J. biol. Chem. **277**: 30445-30453, 2002.

Li, C., Qiao, Z., Qi, W., Wang, Q., Yuan, Y., Yang, X., Tang, Y., Mei, B., Lv, Y., Zhao, H., Xiao, H., Song, R.: Genome-wide characterization of *cis*-acting DNA targets reveals the transcriptional regulatory framework of *Opaque2* in maize. - Plant Cell **27**: 532-545, 2015.

Lopato, S., Forstner, C., Kalyna, M., Hilscher, J., Langhammer, U., Indrapichate, K., Lorkovic, Z.J., Barta, A.: Network of interactions of a novel plant-specific arg/ser-richpProtein, atRSZ33, with atSC35-like splicing factors. - J.biol. Chem. **277**: 39989-39998, 2002.

Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., Misteli, T.: Regulation of alternative splicing by histone modifications. - Science **327**: 996-1000., 2010.

Luco, R.F., Allo, M., Schor, I.E., Kornblihtt, A.R., Misteli, T.: Epigenetic in alternative pre-mRNA splicing. - Cell **144**: 16-26, 2011.

Maor, G.L., Yearim, A., Ast, G.: The alternative role of DNA methylation in splicing regulation. - Trends Genet. **31**: 274-280, 2015.

Motto, M., Maddaloni, M., Ponziani, G., Brembilla, M., Marotta., R., Di Fonzo, N., Soave, C., Thompson, R., Salamini, F.: Molecular cloning of the *o2-m5* allele of *Zea mays* using transposon marking. - Mol. gen. Genet. **212**: 488-494, 1988.

Pirona, R., Hartings, H., Lauria, M., Rossi, V., Motto.: Genetic control of endosperm development and of storage products accumulation in maize seeds. - Maydica **50**: 515-530, **2005**.

Schmidt, R.J., Burr, F.A., Burr, B.: Transposon tagging and molecular analysis of the maize regulatory locus *opaque-2*. - Science **238**: 960-963, 1987.

Schmidt, R.J., Burr, F.A., Aukerman, M.J., Burr, B.: Maize regulatory gene *opaque-2* encodes a protein with a ‘‘leucine-zipper’’ motif that binds to zein DNA. - Proc. nat. Acad. Sci. USA **87**: 46-50, 1990.

Shang, X., Cao, Y., Ma, L.: Alternative splicing in plant genes: a means of regulating the environmental fitness of plants. - Int. J. mol. Sci. **18**: 432, 2017.

Song, R., Llaca. V., Linton. E., Messing, J.: Sequence, regulation, and evolution of the maize 22-kD *α zein* gene family. - Genome Res. **11**: 1817-1825, 2001.

Wang, B.B., Brendel, V.: The ASRG database: identification and survey of *Arabidopsis thaliana* genes involved in pre-mRNA splicing. - Genome Biol. **5**: R102, 2004.

Xiao, X., Wang, Z., Jang. M., Burge, C.B.: Coevolutionary networks of splicing cis-regulatory elements. - Proc. nat. Acad. Sci. USA **47**: 18583-18588, 2007.

Yearim, A., Gelfman, S., Shayevitch, R., Melcer, S., Glaich, O., Mallm, J.P., Nissim-Rafinia, M., Cohen, A.H.S, Rippe, K., Meshorer, E., Ast, G.: HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. - Cell Rep. **10**: 1-13, 2015.